



Williamson, J. R., and Sundén, D. (2016) Deep cover HCI: the ethics of covert research. *Interactions*, 23(3), pp. 45-49.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

© 2016 ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Interactions*, 23(3), pp. 45-49, <http://dx.doi.org/10.1145/2897941>.

<http://eprints.gla.ac.uk/129756/>

Deposited on: 1 December 2016

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Deep Cover HCI: The Ethics of Covert Research in HCI

Julie R. Williamson, University of Glasgow

Daniel Sundén, University of Glasgow and Nile HQ Ltd

In a time when many research questions lead us to evaluate “in the wild,” it seems like the next logical step to increase the realism of these evaluations. Studies done without any interference or visible presence from an experimenter could give us an incredibly realistic view of how our technologies and interfaces are used in practice. The participants may not even realize it is an experiment at all. This would provide an ideal setting for evaluating interaction in the wild, creating not just ecological validity but ecological reality.

At this point, the obvious questions arise about research ethics: What about informed consent? What about data anonymity? However, there are clear guidelines for conducting such research if we look to other disciplines. Sociology, in particular, provides widely accepted guidelines for handling situations where informed consent may not be practical or would disrupt the phenomenon in question. Although covert research has long been a debated area of sociology, there are clear motivations for this technique.

The human-computer interaction (HCI) community has arguably grown out of borrowing and extending (for better or worse) methodologies and theories from other disciplines. Why not the tradition of covert methods, applied to fit the needs, values, and ethics of this community?

Deep Cover HCI is our approach to covertly researching users’ naturalistic responses to interactive technologies. We stage research prototypes as “products” available in the real world, evaluating through interventions in public spaces. After deployment, we maintain cover by avoiding all further intervention or disturbance to the experimental setting. We do not gather qualitative data from users, and we do not actively curate any part of their experience. The resulting data represents users’ uninfluenced and naturalistic responses to the technology being evaluated.

Deep Cover HCI

There are clear motivations to conduct covert research in HCI. The setting of an evaluation or the research questions being addressed can make traditional techniques inappropriate or unhelpful for completing the research.

In many settings, gathering informed consent from every possible participant is impractical and disruptive. For example, when researchers stage evaluations as part of public events, collecting consent forms and qualitative

data through questionnaires heavily interferes with the experience as well as the natural behavior of the participants, possibly skewing any data collected. Evaluations concerned with the usability of prompts without guidance, the attractiveness of displays, and “walk up” experience would be difficult to study meaningfully in a lab. Additionally, data collected in a lab would be influenced in unknown ways as a result of overt observation. Covert research is already being done in the HCI community, but we hope we can open up the discussion about covert methods and address some of the key ethical considerations.

Deep Cover HCI is an intervention-based approach to evaluating technology in public spaces, where the technology in question is staged such that passers-by may not be aware they are part of an experimental setting. Creating an intervention puts the researcher into a special position, where he or she is aware of the intervention and its purpose but passers-by are not. This dynamic creates a form of deep cover, where a researcher is actively influencing the intervention setting as a covert outsider. Maintaining cover in this context means maintaining the secrecy of the experimental purpose.

The key components of our approach consist of:

- Blurring the lines between experimental settings and real-world settings through *evaluation staging*
- *No experimenter intervention* or visible presence after the initial intervention
- Analysis based on multiple streams of *observable data only*
- *No explicit consent* gathered from participants at any point during the evaluation.

Experimental Staging

A key component of Deep Cover HCI is staging experiments such that participants are not aware of the evaluation or the manipulation of experimental variables. Often this means that installations must borrow their heuristics from the professional realm, resulting in installations and experiences with a high degree of fidelity, often indistinguishable from commercial productions. Disguising a research experiment as a commercial product is just one way to hide the purpose of an evaluation.

When staging covert evaluations of technology, there are also practical issues to consider, such as site access, permissions, and logistics. Additionally, deployment hardware must be sophisticated enough to run without an experimenter present and must support an experience that can be completed

without guidance or training. Not all interfaces or technologies are appropriate or make sense for deployment in such a setting.

It is important to recognize the limited settings where staging a covert experiment is ethical. In general, the only ethical settings for completing this work are public and quasi-public settings. Additionally, these settings must be places where participants would reasonably expect to be observed at any given time.

Non-Intervention

In order to maintain cover for the duration of an evaluation, the experimenter must not intervene unless absolutely necessary for health and safety. The presence of an experimenter not only disrupts the staging of the evaluation, it also has unknown and possibly significant effects on the observational data collected.

We do not feel that collecting qualitative data from passers-by is appropriate or useful. First, this would require that an experimenter be present in the deployment space, which might deter potential users or have other unintended effects on observed behaviors. Second, we question the value such data brings, especially given the cost of “breaking cover” in the context of the experiment. We also question the ethics of such an approach if this data were to be collected covertly. If users must be approached at a distance from the installation in order to avoid breaking cover, at what distance can this be achieved practically and ethically? It would be difficult to complete such covert questioning successfully without creating suspicion, making users uncomfortable, or breaking cover.

Naturalistic Observation

Limiting data to that which is observable is key to ethical data collection during covert evaluation. This restriction satisfies ethical guidelines and requires no intervention by the experimenter. For example, technology-supported observation techniques like CCTV analysis gather observable data at scale. Video data can be collected as a constant stream of input from a variety of sources to support both qualitative and quantitative analysis.

A key data source for Deep Cover HCI is behavioral maps generated from video data. Behavioral maps visualize flows of traffic, areas where passers-by crowd, and how passers-by use the space. Behavioral maps support analysis of both interacting and non-interacting users and present data in a completely anonymized format. Video segments can also be used for fine-grain analysis to explore interactions in greater detail.

Detailed on-device logging can provide a view into how users interact with the technology in question and can be tailored to the specific research questions and hardware being used. Such logs provide a fine-grain portrayal of on-device interaction while maintaining the anonymity of users.

Bringing together these data sources, researchers can complete rigorous qualitative and quantitative analysis without the need to intervene during evaluation.

Unwitting Participants

Deep Cover HCI focuses on creating the most realistic evaluations of technology, exploring how users might interact with an interface as if it were part of their everyday life. If an installation is staged well and data collection planned appropriately, participants should not be aware they are participating in an experiment and would thus exhibit uninfluenced responses to and interactions with the technology in question.

If the appropriate constraints are applied to the physical setting where deployments are staged, the hardware/technology evaluated, the data collected, and the role of the experimenter, evaluations can proceed without the need to obtain informed consent from participants. The motivations for the restrictions described above are grounded in ethical guidelines compiled from a variety of authorities.

Ethics

The ethics of completing covert research are the most important part of the broader discussion of in-the-wild studies, particularly when addressing informed consent. The ACM Code of Ethics and Professional Conduct [1] describes our “moral imperatives,” but consent does not feature except in the context of respecting the privacy of others. The Institute for Electrical and Electronics Engineers (IEEE) Code of Ethics [2] is even more laconic on the concept of consent. These codes of ethics do not tackle the specific needs of in-the-wild research. However, there are guidelines for completing in-the-wild and covert research from a multitude of ethical authorities in the social sciences and humanities. These authorities vary in their attitudes toward covert research, but all agree there are times when such research is necessary and that special precautions must be taken.

The Economic and Social Research Council (ESRC), the main source of social science research funding in the U.K., provides a detailed rationale for contexts where covert research is necessary. The ESRC Framework for Research Ethics states: “Informed consent may be impracticable or meaningless in some research, such as research on crowd behaviour, or may be contrary to the research design, as is sometimes the case in psychological

experiments where consent would compromise the objective of the research . . . Covert research may be undertaken when it may provide unique forms of evidence or where overt observation might alter the phenomenon being studied” [3].

The American Sociological Society (ASA) gives a detailed description of settings or contexts where covert research may be appropriate. The ASA states: “Sociologists may conduct research in public places or use publicly available information about individuals (e.g., naturalistic observations in public places, analysis of public records, or archival research) without obtaining consent” [4]. Collecting data based on naturalistic observations in public spaces without consent is a commonly used approach, and generally agreed to be an ethical technique. Completing observational research in public settings where people may expect to be observed does not violate privacy. However, concepts of public/private need to be discussed, as we will see in the guidelines from the European Commission (EC).

The EC, the main funding body for European research, recently completed its Framework Programme 7. This program generated a large amount of documentation on ethics for research in the social sciences and humanities (which at the time of writing, the current program, Horizon 2020, has yet to establish). The Guidance Note for Researchers and Evaluators of Social Sciences and Humanities Research states: “For example in ‘covert research,’ researchers should take into account the meanings of public and private in the contexts they are studying. Covert observation should only proceed if researchers can demonstrate clear benefits of the research, when no other research approach seems possible and when it is reasonably certain that no one will be harmed or suffer as a result of the observation” [5]. This is the most complete guidance in that it brings together the appropriateness of covert research with respect to setting. These guidelines also state: “Another area of ethical concern pertains to the observational research that is central to much sociopsychological research. Observational approaches can vary (focused, participant, invasive/intrusive, visible, covert/overt; recorded rigorously using audio/visual methods or handwritten notes compiled after the event). Researchers should ask themselves several questions concerning the research setting (e.g., is it public or private?), the behavior under scrutiny (in a public or private setting), the way data is collected (recorded or not), and whether or not the protection of participants is ensured.”

In reviewing each of these guidelines, we highlight three key questions to extend basic ethical guidelines that must be addressed to determine if covert research is appropriate:

- Is covert research the only way this data could be collected? For example, is consent impractical, or would consent disrupt the phenomenon being observed?
- Is the setting one where people might reasonably expect to be observed? If not, then covert research may present serious ethical issues.
- What kind of data will be collected, and will the research results maintain the anonymity of those generating the data?

Conclusion

Deep Cover HCI is our approach to covert research in public spaces, restricting our deployments and data collection to complete ethical evaluations. We propose that research completed with high-fidelity prototypes in public spaces, where data is restricted to observable data only, results in the most realistic usage data.

One of the most debated issues in Deep Cover HCI is the purposeful exclusion of qualitative data collection in order to maintain cover during evaluation. Currently, it is difficult to understand what effects the presence of an experimenter may have on the observational data collected. For example, does the behavior of the experimenter deter others from approaching the display? Does the data collected in this way give an unbiased view into user opinions? Until these questions can be answered, we would argue that collecting qualitative data at a deployment site creates unknown bias in observational data.

Because of the significant effort that Deep Cover HCI makes to anonymize data, making data open for review and secondary analysis should be promoted. Making data openly available is becoming a priority for many research councils. For example, the EC Horizon2020 program recently started its Open Data Pilot, with similar initiatives developing in other funding bodies. Open data is also important because it adds transparency to analysis techniques and allows for the open critique of data practice and analysis. Making data publicly available also brings up the question of who owns the data that is generated through observational studies. For example, if a user becomes aware that they generated data in a publicly available dataset, do they have a right to ask for it to be removed? Would this even make sense practically or ethically from a researcher's perspective?

One issue with Deep Cover HCI is the absence of the "reveal" moment, where unwitting participants are made aware of the experiment and its purpose. Although the moment of breaking cover is important in traditional covert research, it's unclear how or when this should be achieved during a long-term covert evaluation.

There is still a need to address more widely the issues and ethics of in-the-wild and covert research in HCI. Covert evaluations are already being completed; issues of ethics, rigor, and methodology are still variable and, importantly, debatable.

Acknowledgments

We would like to thank all of the unwitting participants who have fed our curiosity and enabled us to do the work that intrigues us and expands our understanding of technology in public spaces. This work was supported by the EPSRC SIPS project (publicinteraction.co.uk).

Endnotes

1. Association for Computing Machinery Code of Ethics and Professional Conduct; <http://www.acm.org/about/code-of-ethics>
2. The Institute of Electrical and Electronics Engineers Code of Ethics; <http://www.ieee.org/about/corporate/governance/p7-8.html>
3. ESRC Framework for Research Ethics; <http://www.esrc.ac.uk/about-esrc/information/research-ethics.aspx>
4. American Sociological Association Code of Ethics; <http://www.asanet.org/images/asa/docs/pdf/CodeofEthics.pdf>
5. European Commission Framework Programme 7 Guidance Note for Researchers and Evaluators of Social Sciences and Humanities Research; http://ec.europa.eu/research/participants/data/ref/fp7/89867/social-sciences-humanities_en.pdf

Julie R. Williamson is a lecturer in human-computer interaction at the University of Glasgow. Her research explores how playful interaction in public spaces plays a role in place-making and urban experience. She works with artists, designers, urban theorists, and computing scientists to create, deploy, and evaluate urban interventions. julie.williamson@glasgow.ac.uk

Daniel Sundén is a designer working in convergence between product and user experience and has produced internationally recognized work for major international clients as well as academic research. He has expertise producing physical and digital prototypes as well as experience working with spherical displays for public spaces. daniel@nilehq.com

Insights

- Many research questions can be answered only using covert methods.

- By applying the right constraints to evaluations, covert research can be completed ethically.